PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
H04Q 7/24, 3/00

A1

(11) International Publication Number: WO 97/04611

(43) International Publication Date: 6 February 1997 (06.02.97)

(21) International Application Number: PCT/SE96/00905

(22) International Filing Date: 3 July 1996 (03.07.96)

(30) Priority Data:
08/506,119 24 July 1995 (24.07.95) US

(71) Applicant: TELEFONAKTIEBOLAGET LM ERICSSON (publ) [SE/SE]; S-126 25 Stockholm (SE).

(72) Inventor: LIU, George; Trondheimsgatan 4, S-164 30 Kista (SE).

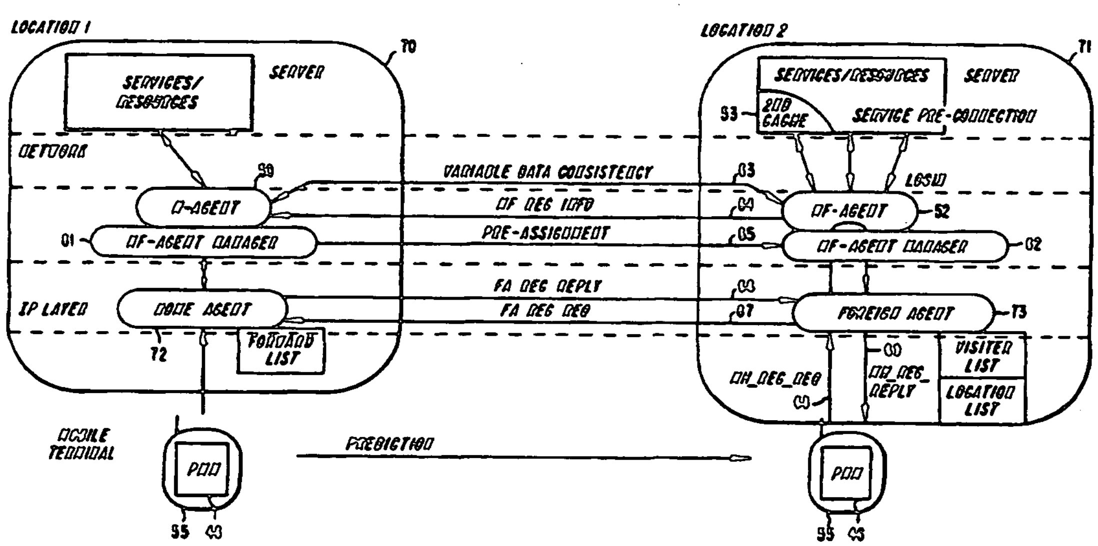
(74) Agents: BOHLIN, Björn et al.; Telefonaktiebolaget LM Ericsson, Patent and Trademark Dept., S-126 25 Stockholm (SE).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: DISTRIBUTING NETWORK SERVICES AND RESOURCES IN A MOBILE COMMUNICATIONS NETWORK



(57) Abstract

A full mobility data network architecture and method supports global wireless mobile data accessing. A Mobile Distributed System Platform, a mobility agent and a mobile-floating agent are provided for supporting service and resource mobility, and for distributing network services and resources closer to mobile users. In one aspect of the invention, a predictive mobility management algorithm determines where a mobile user or terminal is likely to be. Mobile-floating agents are then established at these locations to permit pre-connection and pre-arrangement of network services and resources for use by the user upon arrival. This allows mobile networks to more intelligently and dynamically provide services to mobile users. The Mobile-Floating Agent scheme can be used, for example, to build a mobile virtual distributed home location register (HLR) and visitor location register (VLR) in personal communication networks (PCNs), in order to reduce the call set up time. It can also be used for "soft data structure handover" for mobile computing.

DISTRIBUTING NETWORK SERVICES AND RESOURCES IN A MOBILE COMMUNICATIONS NETWORK

BACKGROUND

The present invention relates to methods and apparatus for supporting data and service mobility to users of mobile networks.

With the freedom gained by increased access of external data through mobile user networks, productivity of computers and their users will be increased. Productivity will also be increased when mobile systems' users can efficiently access their data in addition to efficiently managing their voice calls. Generally mobile computer users require more access to network resources, such as files and databases, than do mobile telephone users. As a result, how to efficiently provide mobile data access is an important question.

Previous research in this area has resulted in a number of proposals that address network layer mobility support for mobile computing, including Mobile Internet Protocol (IP), sending packet data over mobile radio channels, and network layer host migration transparency. Although it is possible for a mobile computer to access data through the use of mobile networks, inefficient mobility support can result in performance problems such as poor throughput. Current cellular telephone networks are not efficient for wireless data access because they do not support data and service mobility. While users and their terminals are mobile, the fact that their data is configured statically in the system remains a problem that inhibits efficient data access.

In today's systems, databases, such as Home Location Registers (HLRs), are designed and configured centrally. Central databases are inefficient for a large number of mobile users because they do not support

previous location. It also provides "soft data structure handoff" capability.

In accordance with one aspect of the invention, network services and resources are distributed to a mobile user in a mobile communication system by providing the mobile user with a mobility (M)-agent executing on a home fixed host or router. It is then determined that the mobile user is or will be travelling to a destination that is outside a service area of the home fixed host or router, and a pre-assignment request is sent from the M-agent to at least one mobile floating (MF)-agent manager executing on a corresponding one of a like number of remote fixed hosts or routers located at the destination. Each MF-agent may include a set of processes, executing on the corresponding remote fixed host or router, for communicating and connecting with local resources and for managing a variable replicated secondary data cache on behalf of the M-agent. The M-agent may include a set of processes, executing on the home host or router, for communicating with each MF-agent. A mobile floating (MF)-agent is then established for use by the mobile user at each of the remote fixed hosts or routers, and the M-agent is used to send data or service information from the service area of the home fixed host or router to the MF-agent at each of the remote fixed hosts or routers. In this way, services and/or data may be pre-connected/pre-arranged at the mobile user's destination.

In accordance with another aspect of the invention, the step of establishing the MF-agent comprises, at each of the remote fixed hosts or routers, determining whether any preexisting MF-agent exists at the remote fixed host or router, and if not, then creating the MF-agent for use by the mobile user at the remote fixed host or router. Otherwise, the preexisting MF-agent is

In an alternative embodiment of the invention, the step of identifying the corresponding remote fixed hosts or routers located at the destination comprises the steps of: sending a message from the mobile user to the M-agent, wherein the message designates the destination; and using the M-agent to identify the MF-agent manager that is executing on a remote fixed host or router that is located at the destination. In another aspect of the invention, the message may further include a designated time that the mobile user will be at the destination. In this case, the additional step of transferring data from the M-agent to a secondary cache of the MF-agent is performed, wherein the time of the data transfer has a predetermined relationship with the designated time. For example, the predetermined relationship may require transfer of the data prior to the designated time.

In yet another alternative embodiment of the invention, the step of identifying the corresponding remote fixed hosts or routers located at the destination comprises the steps of: (1) determining a mobility density m, wherein m is a number of cells that have been passed by the mobile user during time τ_m; (2) defining a circularly shaped geographic location centered at a current location of the mobile user and having a radius $d=int(h*m*\tau_m)$, wherein d is a service distance and h is a hierarchic factor defined by the number of cells serviced by one MF-agent manager; (3) predicting a movement track (MT) or movement circle (MC) of the mobile user; and (4) identifying the remote fixed hosts or routers that are located on the predicted MT or MC within the circularly shaped geographic location.

In still other aspects of the invention, the disclosed techniques may be applied in wireless local area networks (LANs) as well as in cellular communications systems. Further, a mobile-application

vs the Average Mobility Rate for the Radius-d Assignment Scheme;

FIG. 16 illustrates the Performance gain vs. Mobility density for the MT/MC/d assignment method;

FIG. 17 illustrates the percentage of Latency reduction vs. Mobility Density for the MT/Mc/d Assignment method;

FIG. 18 shows the general Architecture of a Wireless LAN;

FIG. 19 is an example of a useful testbed for use with the present invention;

FIG. 20 depicts the protocol architecture of a wireless LAN;

FIG. 21 shows an example of the Mobile Floating Agent on the MSR Protocol Architecture;

FIG. 22 illustrates the relationship between M-agents and T-agents.

FIG. 23 is an example of Mobile Floating (MF)-Agent implementation;

FIG. 24 is an example of a cellular architecture; and

FIGS. 25A and 25B show a cellular communication system implementation and an example of MF-Agent (MFA) implementation on a cellular communication system.

DETAILED DESCRIPTION

The various features of the invention will now be described with respect to the figures, in which like parts are identified with the same reference characters.

A Comprehensive Mobility Architecture

With the proliferation of mobile communication into everyday life, work environments have become decentralized. As a result, users can now take work

outlined above, service and resource mobility provided by these networks is becoming increasingly important for providing efficient mobility management. Service mobility is the mobility of various services (logic/data) in the underlying access 12 and backbone 14 networks. Resource mobility is the mobility of the resources, such as system data/programs, user data, user programs, etc., in the underlying network. Both service and resource mobility must be adequately provided in order to meet the quality of service requirements of the mobile users.

In order to meet the increasing demands that user mobility is placing on these networks, conventional mobility management capabilities must be further extended to manage service and resource mobility. The importance of these two additional types of mobility is significant for efficient mobility management support.

To efficiently support mobility, the user agent model as outlined in Lennart Söderberg, "Evolving an Intelligent Architecture for Personal Telecommunication," ERICSSON Review, No. 4, 1993, hereby incorporated by reference, is introduced. Referring now to FIG. 2, each user 21 and terminal 22 is represented in the network by corresponding agents 24 and 25 respectively. These agents 24 and 25 contain all service logic and service data related to the user 21 or terminal 22, and control all communication sessions of the user 21 or terminal 22. This model provides the basis for providing service and resource mobility in accordance with one aspect of the invention.

Referring now to FIG. 3, mobile terminal software 39 in accordance with one aspect of the invention is shown. Both a Mobile-Applications Programming Interface (API) 31 and a Mobile Floating (MF)-agent 38 are provided to cope with the varying bandwidth and

communications with a mobile system or other terminals while it is changing location. However, one may typically expect to encounter different types of connectivities (different radio/IR bandwidths) and services/resources (e.g., servers, printers, programs, etc.) at different locations in a wireless network environment. In order to ensure efficient management of location-sensitive information, applications must determine the characteristics of the communication channel and services provided by the networks and/or be notified of changes. Therefore, location-sensitive information, identifying the services or resources (including hardware and software resources, network connectability and types of communication protocol available, etc.) provided by the systems or networks at a defined location (i.e., geographical area), must be efficiently managed.

The LSIM 47 in the MDSP 45 is designed to manage the location-sensitive information and map the information to the different services offered by the mobile infrastructure at different geographical locations. Furthermore, the LSIM 47 is also responsible for informing both the applications 30 and their supporting agents 38 about any change of location of the mobile terminal 22 in addition to providing dynamical service connections. For example, suppose that a network, with a distributed file system, consists of several servers distributed in different geographic areas. When a mobile terminal moves from a location near server A to a location near server B, the LSIM 47 should inform both the server B and a cache manager in the mobile terminal, that server B is the nearest file server, should a file be needed.

In one embodiment of the invention, the most likely destination of a user is determined through the

and resources are not bound to the underlying network. Therefore, the M-agent 50 and MF-agent 52 are free to follow the mobile users. By using predictive mobility management to predict where the user will be, as described in U.S. Patent Application No. 08/329,608, filed on October 26, 1994, incorporated by reference, the MF-agent 52 pre-connects services, pre-arranges the secondary cache 53 and prefetches data from the home user cache 51 to be placed in the secondary cache 53, in a fashion similar to that in which a travel agency would pre-arrange a hotel room or other services for a user when that user is travelling.

Mobile Floating Agent Protocol

The MF-agent 52 is assumed to have basic mobility support from the network layer, via a protocol such as Mobile-IP, otherwise known as IP mobility support.

A preferred embodiment of the MF-agent pre-assignment protocol is depicted in FIGS. 6 and 7. The MF-agent manager 61, 62 provides a common base which supports MF-agent 52 creation and assignment (i.e., establishment). The M-agent 50 is a representative of the user 21 in the network and is responsible in part for creating, deleting and managing the MF-agents on behalf of mobile users. An M-agent 50 requests creation or assignment of MF-agents 52. As shown in FIG. 7 a mobile terminal 55 sends an MF-agent assignment request to its M-agent 50, in the local network, with an address of a new location it is travelling to (701). The new location may be one that has been explicitly provided by the user 21, or it may be one predicted by the PMM functions 46. The assignment request is a request to establish (i.e., alternatively create or pre-assign) an MF-agent 52 at the location that the mobile terminal 55

terminal moving from New York City to Europe), then the data consistency link 63 can have a low priority requiring less frequent updating of the secondary cache 53. If the time needed to reach the new location is shorter, (for example, driving within New York City), then the data consistency link 63 would be given a higher priority requirement for more frequent updating. Data consistency and caching methods are described in more detail below.

Referring now to FIG. 8, when the mobile terminal 55 reaches the new location, it registers with the MF-agent 52 that has been created or assigned for it there (801). This is accomplished by sending an MF-agent registration request 68 to the F-Agent 73 at the new location to begin the registration process. The F-agent 73 checks to see if there is a corresponding MF-agent 52 for the mobile terminal 55 (802). If there is an MF-agent 52, the F-agent 73 confirms this and activates the MF-agent 52 (804). The F-agent 73 then links the mobile terminal 55 to the MF-agent 52 (805). In accordance with another aspect of the invention, the MF-agent now performs as an acting M-agent (AM-agent) for the mobile terminal 55, performing the same function as an M-agent at the new location. It should be noted that the M-agent 50 at the home location will always be the M-agent and is responsible for control of all communication sessions. The M-agent 50 also maintains data consistency between the home resources (data/files) and its pre-assigned MF-agents and its AM-agent on behalf of its user, the M-agent 50 being able to maintain more than one data consistency link 63. (An AM-agent may also maintain multiple data consistency links 63 with other MF-agents 52. This arrangement is shown in FIG. 10 and described in more detail below). Once a mobile terminal 55 moves to yet another new location and a new

parameter provides a priority index for shared resources (e.g., disk space for secondary cache, memory, etc.) with other MF-agents 52 at the same location. If the resources at this location have to be reclaimed, then one or more MF-agents 52 having the highest LRU parameter are chosen as victims and destroyed in order to free the needed resources.

Mobility-aware Dynamic Caching and Prefetching Method

According to another aspect of the invention, a method of hierarchic Mobility-Aware Dynamic (MAD) cache management is provided, for dynamically managing and updating the secondary cache of the MF-agent 52. The following subsections describe the basic principles of the hierarchic MAD caching, Dynamical Caching Consistency (DCC) and Mobility-aware Caching Management.

Hierarchical MAD Caching Consistency

FIG. 9 illustrates a MAD Caching Scheme in accordance with one embodiment of the invention. The MAD Caching Scheme is designed preferably in a hierarchy with two classes of caches: a primary cache 91 implemented at the terminal and a secondary cache 92 that is managed by the MF-agent.

Two classes of cache consistency methods are employed to maintain data consistency. A first class utilizes a dynamic cache consistency (DCC) method to maintain data consistency between a server and the MF-agents 52, including any AM-agents. A second class includes a mobility-aware cache coherence method, wherein the MF-agent 52 keeps track of any items cached by its mobile user and is responsible for broadcasting an invalidation report if any of the items are changed.

for the mobile terminal 55 in this cell). Thus, the total number of invalidation reports broadcasted at each cell can be reduced. FIG. 11 shows an example of invalidation report broadcast areas for different mobile terminals. These areas may overlap, change over time and move when the mobile terminals change mobility behavior during each time period τ_m.

A second advantage of using MF-agents 52 for broadcast invalidation reporting is that the invalidation database is only replicated at the cell or station within each user's dynamic location area. This can reduce the update overhead for fewer mobile users because their corresponding dynamic location areas will be small, for example 1101 or 1104. One can think of this as an information radius of the mobile terminal, which decreases as the certainty of the user's location increases.

An additional advantage is that with the support of the MF-agents 52 in the fixed network, individualized dynamic invalidation reports can be defined for each mobile terminal 55 according to the terminal's mobility behavior or the cache consistency required during each time period τ_m.

Furthermore, through the use of the MF-agents 52, a replicated database is created for each cell or MSR, corresponding to each individual mobile terminal 55. The database is dynamically replicated about the mobile user's location and changes according to the user's mobility. By utilizing the PMM functions of the MDSP and the MF-agent protocol, whenever a user moves to another location, the user will always find the data that is needed replicated at that location.

The secondary caches of the MF-agents 52 can be constructed so that they perform the same functions as a Page-answer Database, detailed in N. Kamel & R. King,

type 2 DCC 1002, 1022. The type 2 DCCs 1002, 1022 are preferably lower priority links used to update the caches of associated MF-agents. The number of MF-agents 1003, ..., 1009, 1020, 1030 and their relation to the AM-agent 1006 and to the MA 93 as shown in FIG. 10A is just an example. In practice, the MF-agents 1003, ..., 1009, 1020, 1030 can exist in any number, and can be distributed in different patterns depending on the MF-agent assignment method used, such as Movement Circle (MC) and Movement Track (MT) patterns.

In one preferred embodiment, the type 1 DCC 1001 uses a "call-back" consistency policy. The M-Agent 93 is responsible for keeping track of the cache status information of its current AM-agent 1006. To avoid frequent changes in the association with different agents as the terminal moves back and forth, the old AM-agent 1006 (FIG. 10A; also MFA_0 1016 in FIG. 10B) is preferably used to forward the type 1 DCC 1001 to the new AM-agent 1013. The association between an M-Agent 93 and the old AM-agent 1006 will last for a period of time τ_{dl} even when the mobile terminal 55 has moved to another location. The value of τ_{dl} is given by

$$\tau_{dl} = \frac{h\alpha_{dl}}{m}$$

where α_{dl} is a delay factor, h is a hierarchy factor and m is the mobility density of a user during time period τ_m.

The new AM-agent 1003 informs the M-Agent 93 to establish the type 1 DCC 1101 association with it after a time period τ_{dl}. The M-Agent 93 is only allowed to have a type 1 DCC association with one of the MF-agents associated with each mobile user; all other type 1 DCCs 1001 associated with the old AM-agents 1006 will be canceled after the M-agent forms an association with a

AM-agent 1006) before 1/m, in other words, the time needed for a mobile user with mobility density m to move from one cell to another. During the delay time τ_{α} there may be several updates. Only the latest update is group-cast to the MF-agents in the group.

MF-Agent Assignment

The above sections have outlined the MF-agent protocols and DCC methods. However, in order to use an MF-agent it is also necessary to identify the corresponding fixed hosts or routers located at the destination where the MF-agents need to be created or assigned.

Radius-d Assignment Method

Consider a geographic area covered by cells (such as those defined by a cellular network), with each cell being serviced by an MSR or fixed host. Let λ be the average move rate of a mobile user, where average move rate is defined by the average number of new MSRs which have been passed by the mobile user during a relatively long unit time, such as a day, month or year. Let m be the mobility density factor of the mobile user, where mobility density factor is defined as the number of new cells that have been passed by the mobile terminal 55 during time τ_m. Let s denote service rate, which is defined by the number of MF-agents that service each unit movement of the mobile users. Then:

$$s = \frac{d}{h \cdot m \cdot \tau_m}$$

where d is the service distance, h is a hierarchic factor preferably defined by the number cells serviced

generated in the fixed network due to a large number of MF-agents being assigned within a circle having a large radius d. Another problem is that some of the MF-agents assigned in the circle may never be used. This is an inefficient use of resources. A better method of assignment is to combine the MF-agent assignment with the predictive mobility management (PMM), as fully described in previously mentioned co-pending U.S. Patent Application No. 08/329,608, filed on October 26, 1994, incorporated herein by reference. Instead of assigning MF-agents to every remote fixed host or router within the circle of radius d, this assignment method predicts the most likely movement of the mobile user within the circle. The MF-agents are only assigned to those remote fixed hosts or routers within the circle that are additionally within a predicted Movement Track (MT) or Movement Circle (MC). As an MT or MC can cover a very long physical distance, only the states or cells in the MT or MC within the distance d are assigned each time. This method is referred to throughout this description as the MT/MC/d assignment method. This method provides a more efficient allocation of MF-agents than the bare Radius-d method.

To state this more succinctly, the MT/MC/d assignment method is as follows:

1. Calculate the mobility density m during each time interval τ_m.

2. Define a circle centered at the current location of the mobile terminal having a radius d. (Assume that $\alpha_d = 1$, i.e., 100% confidence that the user is within the circle with radius d during τ_m). The radius d is given by

$$d = \mathrm{int}(hm\tau_m)$$

3. Only assign MF-agents to those MSRs that are located on those portions of the predicted MT or MC that

utilizes only two input parameters: the "future-location" (network address) and the time (t_c) to be expected at that location. However, the time (t_c) can be considered optional. If a time parameter is not specified, the assignment is assumed to have a low priority and will be performed whenever the network is free.

Three MF-agent assignment methods have been described. The Radius-d assignment guarantees a 100% service rate ($s \geq 1$) for any type of random movement by users, but it may generate a lot of background system overhead. By comparison, the point-to-point assignment method may not generate as great an overhead as the Radius-d assignment method does, but it requires explicit "future-location" information in order to provide good service. The MT/MC/d assignment algorithm provides a compromise solution by using the predictive mobility management information to reduce the overhead while avoiding the need for explicit future location information.

Performance Evaluations

Next, the performance of the MF-agent scheme with different assignment methods is evaluated.

Mobile data accesses primarily comprise two important operations, namely, the read and write operations. A read operation reads data from a remote server or database and a write operation writes data back to the server. Therefore, the performances of a read and write operation in the different MF-agent assignment methods are analyzed and evaluated below.

To analyze the performance gain without loss of generality, let the total number of cells be N and also assume that $N >> m \cdot \tau_m$, where $m \cdot \tau_m$ is the total number of

reduction in delay of read and write operations in a system with a PTP assignment is a function of the relative assignment distance, as shown in FIG. 12.

The relative assignment distance D′ is defined as the wired-network latency versus the latency of the wireless network, that is

$$D' = \frac{D\tau_{c3}}{\tau_{c1}}$$

where τ_{c3} is the latency of the fixed network per unit length, while τ_{c1} is the latency of the wireless network.

As stated earlier, it is assumed that the links in the fixed network have a higher bandwidth than the wireless ones, but the longer the distance of the link, the larger the latency will be. This is due to delays (through switches, routers, and nodes) encountered in the link. Therefore, the significance of the relative assignment distance D′ is that the distance of the wired network link is adjusted by the fraction of the wired and the wireless latency. For example, D′ = 1 means that the distance of the assignment is so long that the latency of the wired link ($D'\tau_{c3}$) is the same as that of the wireless one (τ_{c1}).

In FIG. 13, it can be seen that, with the PTP assignment scheme, the percentage of reduced latency increases as the relative assignment distance increases. The delay of read and write operations can be reduced by more than 65% when the relative assignment distance D′ is larger than 0.9. Therefore, the PTP scheme is more suitable for long distance assignments, such as travelling from New York to Europe.

$$G = \frac{L_{w/out} - L_{MF}}{L_{w/out}}$$

$$= \frac{1.4\tau_{c_1} + (4-0.09\tau_{c_1}) \cdot \left[\alpha_D \lambda T \tau_{c_3} \sum_{k=\lambda T}^{\infty} \left[\frac{(\lambda T)^k e^{-\lambda T}}{k!}\right]\right]}{4\left[\tau_{c_1} + \alpha_D \lambda T \tau_{c_3} \sum_{k=\lambda T}^{\infty} \left[\frac{(\lambda T)^k e^{-\lambda T}}{k!}\right]\right]}$$

where $\frac{1}{2\pi} \le \alpha_D \le 1$.

In FIG. 15 the percentage of latency reduction G versus the average movement rate λ is plotted. It is assumed that T=100 minutes $\alpha_D = 0.5$, $\alpha_d = 1$, and $\tau_{c3} = 1$ second.

From FIG. 15 it can be seen that the Radius-d method cuts the latency of read and write operations by more than 60% for users with an average mobility rate larger than 0.2. The higher the mobility of the users, the more the percentage of latency is reduced. This shows that this scheme is very efficient in reducing mobility related latency.

With MT/MC/d Assignment

With the MT/MC/d assignment method, the MF-agents are only assigned to the fixed hosts or router that are located on the predicted MC or MT pattern within circle d. This can greatly reduce the total number of MF-agents assigned and avoid many unnecessary assignments, and thus, reduce the background overhead costs for high mobility density users.

FIG. 16 shows the total performance gain with the MT/MC/d assignment algorithm versus mobility density.

accuracy rates of 0.9, 0.8, 0.7 and 0.6 respectively. It is noted that the percentage of latency reduced is not so sensitive to the average prediction accuracy rate. Even when 40% of the time the prediction is wrong, the latency can still be reduced by more than 65% in the mobility density interval [0.4, 1].

The performance of the MF-agent scheme with different assignment methods, namely Point-to-Point, Radius-d and MC/MT/d have been analyzed and evaluated above. Detailed analyses of these schemes in terms of latency reduction have been presented. The evaluation results show that the MF-agent scheme with the predictive MC/MT/d assignment method can improve overall performance by more than 25% for any mobility density of users and can reduce the access latency for normal mobility density users in the range (0.2, 1) by more than 55%.

Mobile Floating Agent Implementation Issues

FIG. 18 shows a general wireless LAN architecture. The general architecture of a wireless LAN connected by Internet may consist of two basic parts: 1) fixed LANs (e.g. ethernet, ring, etc.); and 2) wireless interfaces, such as Mobile Support Routers with radio and IR.

FIG. 19 is an example of a good testbed for the present invention. The testbed architecture shown in FIG. 19 includes 5 LANS, interconnected by a Gigabit Network. The mobile-IP, as outlined in IETF, "IP Mobility Support," Networking Working Group Internet Draft 11, April 1995, hereby incorporated by reference, is used as the basis for the network layer protocol. The network consists of Walkstations, or mobile

bottleneck. This occurs when mobile terminals **100** access fixed servers **106** through the serial path of the wireless link **108** and wired link **109**. This means that the bandwidth of the access path is limited by the serial bottleneck, that is, the wireless link part **108**. No matter how fast the transmission speed of the fixed network is, the speed of the round-trip path is blocked by the bottleneck. This kind of bottleneck transmission may degrade the speed of the fixed network.

Another problem that can be encountered is Transmission Control Protocol (TCP) timer back-off. The TCP provides support for stream oriented end-to-end communication with reliable delivery. This requires retransmission associated timers to guarantee reliable delivery by retransmission should a time-out occur. The serial bottleneck problem of a wireless link, associated with temporary disconnection, temporary failure, or handover, for example, may cause a TCP timer back-off. For example, if a mobile terminal has an open TCP connection to a fixed host and at the same time the mobile terminal becomes temporarily disconnected from the network, the mobile terminal is unable to retrieve data from the TCP window. As the data is sent to the TCP window on the fixed host, the TCP window fills. Once the window fills, it becomes unable to store any more data. This will cause the TCP to back off and retransmit less and less frequently. After reconnecting, the mobile terminal will have to wait for the next retransmission from the fixed host before it can continue receiving its data.

Most of the above-described problems, as well as the problems discussed in the previous sections of this disclosure, can be solved by the MF-agent manager with its MF-agent. The MF-agent manager and its MF-agent can be implemented at Mobile Support Routers (MSRs), at

predicted by the PMM in the mobile terminal 100 as discussed above), the T-agent will pre-arrange the movement by informing the MF-agent manager 104 in the other MSR 102 to create an MF-agent for its mobile terminal and to manage the handover process.

5. Upon logging in by the mobile terminal 100 at the new area, the MF-agent is active as a T-agent and sends a message back to de-activate the previous T-agent.

FIG. 23 shows an exemplary implementing of the MF-agent 1901 on top of the I-TCP 1903 at an MSR 1905, in accordance with one aspect of the invention.

The Mobile Floating Agent in Cellular Systems

Referring to FIG. 24, a conventional cellular architecture generally comprising a Mobile Switching Center (MSC) 140, base station systems (BSS) 135 and mobile terminals 100. The BSS 135 can be subdivided into different subsystems, such as, the Base Station Controller (BSC) 130 and Base Transceiver Station (BTS) 120. The service area is divided into cells. Each cell is covered (i.e., served) by a BTS 120 operating on a set of radio channels which are different from the ones used in neighboring cells in order to avoid interference. A group of BTSs 120 is controlled by a BSC 130 which also controls such functions as handover and power control. A number of BSCs 130 are served by an MSC 140. The Gateway MSC (GMSC) 140 controls calls to and from the Public Switched Telephone Network (PSTN) 150, Integrated Services Digital Network ISDN 160, Public Data Network (PDN) 170, and the like.

called predictive mobility management has been developed. A Predictive Mobility Management (PMM), as described previously, is used to predict the future location of a mobile user according to the user's movement history patterns.

The combination of the mobile floating agent concepts with the predictive mobility management allow for service and resource pre-arrangement. The data or services are pre-connected and assigned at the new location before the user moves into the new location. As a result, the user can get his/her service or data accessed with virtually the same efficiency as at the previous location.

The present invention has been described by way of example, and modifications and variations of the exemplary embodiments will suggest themselves to skilled artisans in this field without departing from the spirit of the invention. The preferred embodiments are merely illustrative and should not be considered restrictive in any way. The scope of the invention is to be measured by the appended claims, rather than the preceding description, and all variations and equivalents which fall within the range of the claims are intended to be embraced therein.

MF-agent with a foreign agent (FA) located at the remote fixed host or router; and

sending an establishment reply from the MF-agent manager to the M-agent, wherein the establishment reply includes information for enabling the M-agent to establish a data consistency link with the MF-agent.

4. The method according to claim 3, wherein each MF-agent includes a set of processes, executing on the corresponding remote fixed host or router, for communicating and connecting with local resources and for managing a variable replicated secondary data cache on behalf of the M-agent.

5. The method according to claim 3, wherein the M-agent includes a set of processes, executing on the home host or router, for communicating with each MF-agent.

6. The method according to claim 3, further comprising the step of using one of the MF-agents to perform acting mobility (AM)-agent functions in response to the mobile user having logged in at the remote fixed host or router that corresponds to the one of the MF-agents.

7. The method according to claim 6, wherein the step of using the one of the MF-agents to perform AM-agent functions includes the steps of:

determining that the mobile user is or will be travelling to a second destination that is outside a remote service area of the remote fixed host or router that corresponds to the one of the MF-agents;

sending the pre-assignment request from the one of the MF-agents to at least one other MF-agent manager executing on a corresponding other one of at least one remote fixed host or router located at the second

the one of the remote fixed hosts or routers;

selecting one of the identified MF-agents that has a corresponding parameter that indicates that the selected one of the identified MF-agents is the least recently used of all of the identified MF-agents; and

reclaiming resources from the selected one of the identified MF-agents.

10. The method according to claim 3, further comprising the steps of:

sending a registration request from the mobile user to the foreign agent; and

in response to receipt of the registration request, if there is an MF-agent registered for the mobile user, then sending, from the FA to the mobile user, a confirmation and a link for accessing the MF-agent.

11. The method according to claim 1, wherein the step of sending the pre-assignment request from the M-agent to at least one mobile floating (MF)-agent manager executing on the corresponding one of at least one remote fixed host or router located at the destination comprises the steps of:

identifying the corresponding remote fixed hosts or routers located at the destination; and

sending the pre-assignment request from the M-agent to each of the MF-agent managers executing on one of the identified corresponding remote fixed hosts or routers,

wherein identifying the corresponding remote fixed hosts or routers located at the destination comprises the steps of:

(1) determining a mobility density m, wherein m is a number of cells that have been passed by the mobile user during time τ_m;

(2) defining a circularly shaped geographic

designated time.

14. The method according to claim 1, wherein the step of sending the pre-assignment request from the M-agent to at least one mobile floating (MF)-agent manager executing on the corresponding one of at least one remote fixed host or router located at the destination comprises the steps of:

identifying the corresponding remote fixed hosts or routers located at the destination; and

sending the pre-assignment request from the M-agent to each of the MF-agent managers executing on one of the identified corresponding remote fixed hosts or routers,

wherein identifying the corresponding remote fixed hosts or routers located at the destination comprises the steps of:

(1) determining a mobility density m, wherein m is a number of cells that have been passed by the mobile user during time τ_m;

(2) defining a circularly shaped geographic location centered at a current location of the mobile user and having a radius $d=int(h*m*\tau_m)$, wherein d is a service distance and h is a hierarchic factor defined by the number of cells serviced by one MF-agent manager;

(3) predicting a movement track (MT) or movement circle (MC) of the mobile user; and

(4) identifying the remote fixed hosts or routers that are located on the predicted MT or MC within the circularly shaped geographic location.

15. A method for distributing network services and resources to a plurality of mobile users in a mobile communication system in order to ensure that at least one mobile floating (MF)-agent is supporting each mobile user, the method comprising the steps of:

(4) establishing a mobile floating (MF)-agent for use by the mobile user at each of the identified remote fixed hosts or routers.

17. A method for distributing network services and resources to a plurality of mobile users in a mobile communication system in order to ensure that at least one mobile floating (MF)-agent is supporting each mobile user, the method comprising the steps of:

for each one of the mobile users,

(1) determining a mobility density m, wherein m is a number of cells that have been passed by the mobile user during time τ_m;

(2) defining a circularly shaped geographic location centered at a current location of the one of the mobile users and having a radius $d=int(h*m*\tau_m)$, wherein d is a service distance and h is a hierarchic factor defined by the number of cells serviced by one MF-agent manager;

(3) predicting a movement track (MT) or movement circle (MC) of the mobile user;

(4) identifying the remote fixed hosts or routers that are located on the predicted MT or MC within the circularly shaped geographic location; and

(5) establishing a mobile floating (MF)-agent for use by the mobile user at each of the identified remote fixed hosts or routers.

18. A method of distributing network services and resources in a wireless local area network having first and second mobile support routers (MSR) and a mobile terminal, the method comprising the steps of:

in the first MSR, creating a first terminal agent (T-agent) for representing the mobile terminal in the wireless local area network;

means for handing over representation of the mobile terminal from the T-agent to the MF-agent when the mobile terminal arrives in a service area of the second MSR.

23. The apparatus according to claim 22, wherein the MF-agent manager is supported by a mobile internet protocol (IP) layer of the wireless local area network.

24. The apparatus according to claim 22, wherein the first T-agent manages a cache and mobility information of the mobile terminal.

25. The apparatus according to claim 22, further comprising means, responsive to logging in by the mobile terminal at the service area of the second MSR, for causing the MF-agent to act as a second T-agent and to send a message to deactivate the first T-agent.

26. A wireless local area network (LAN), comprising:

at least one mobile terminal;

a mobile support router (MSR) connected to the mobile terminal, the MSR including:

a mobile internet protocol (IP) layer; and

a mobile floating (MF)-agent manager, supported by the mobile IP layer, and having associated mobile floating (MF)-agents; and

at least one fixed host or server connected to the MSR.

27. The wireless LAN according to claim 26, wherein the mobile IP layer includes an indirect transmission control protocol layer for supporting stream oriented end-to-end communication in the wireless LAN.

mobile terminal from the T-agent to the MF-agent when the mobile terminal arrives in a service area of the second base station system; and

the second base station system.

29. The cellular communications system according to claim 28, wherein the MF-agent includes a set of processes, executing on the second base station system, for communicating and connecting with local resources and for managing a variable replicated secondary data cache on behalf of the mobile terminal.

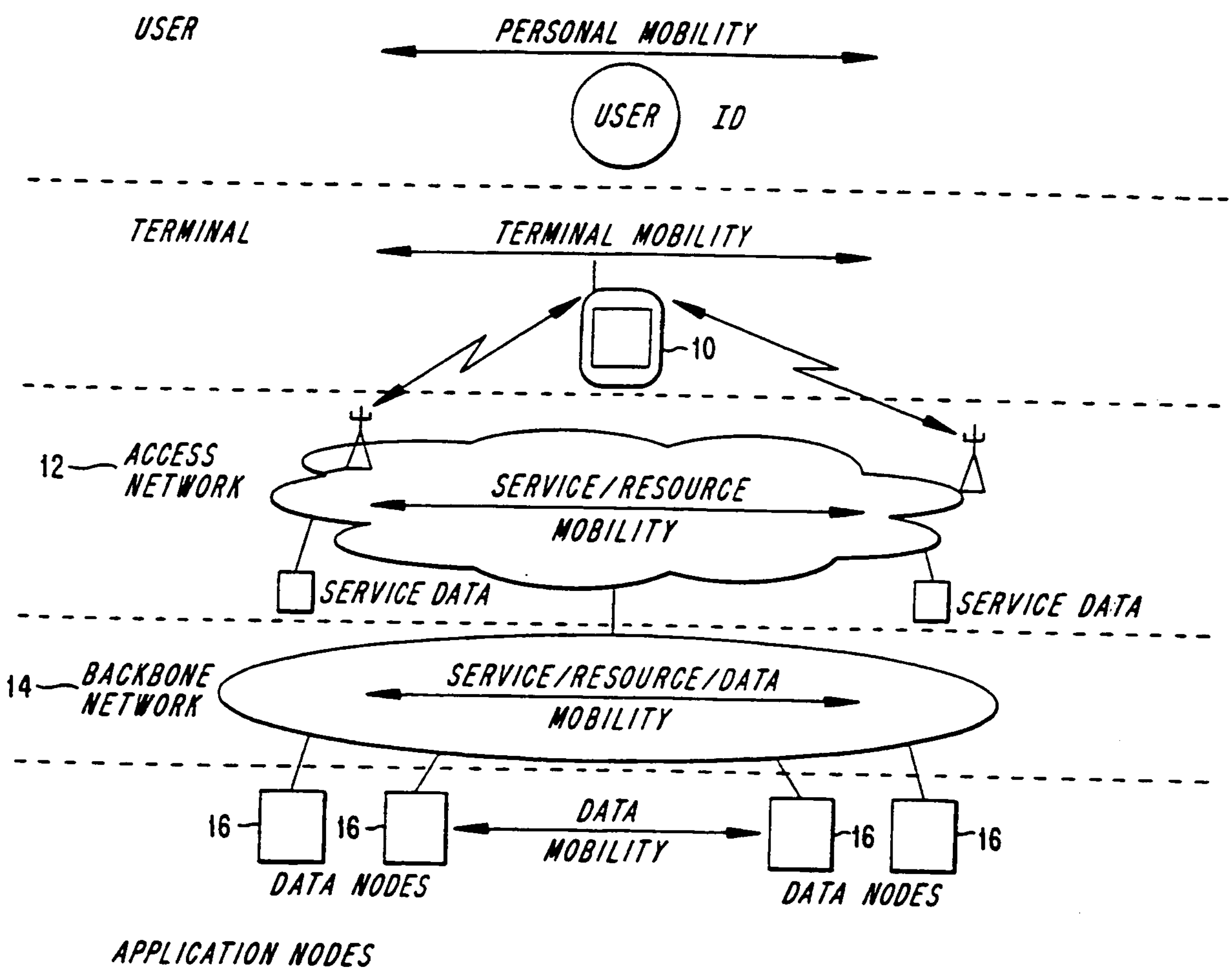
30. The cellular communications system according to claim 28, wherein the MF-agent includes a set of processes, executing on the second base station system, for communicating and connecting with local resources and for managing a variable replicated secondary data cache on behalf of a user of the mobile terminal.

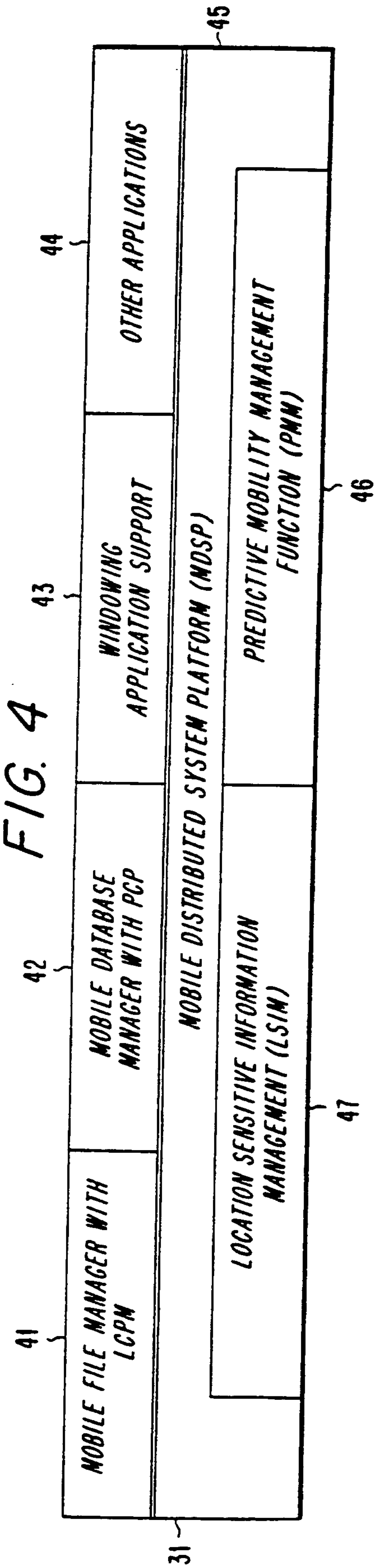
31. The cellular communications system according to claim 28, wherein the mobile terminal includes a mobile distributed system platform for performing location sensitive information management functions and for performing predictive mobility management functions.

32. The cellular communications system according to claim 28, further comprising a gateway router for coupling the mobile switching center to the fixed communications network, wherein the gateway router includes means for creating a second MF-agent for representing the mobile terminal in the cellular communications system.

33. The cellular communications system according to claim 28, further comprising a gateway router for

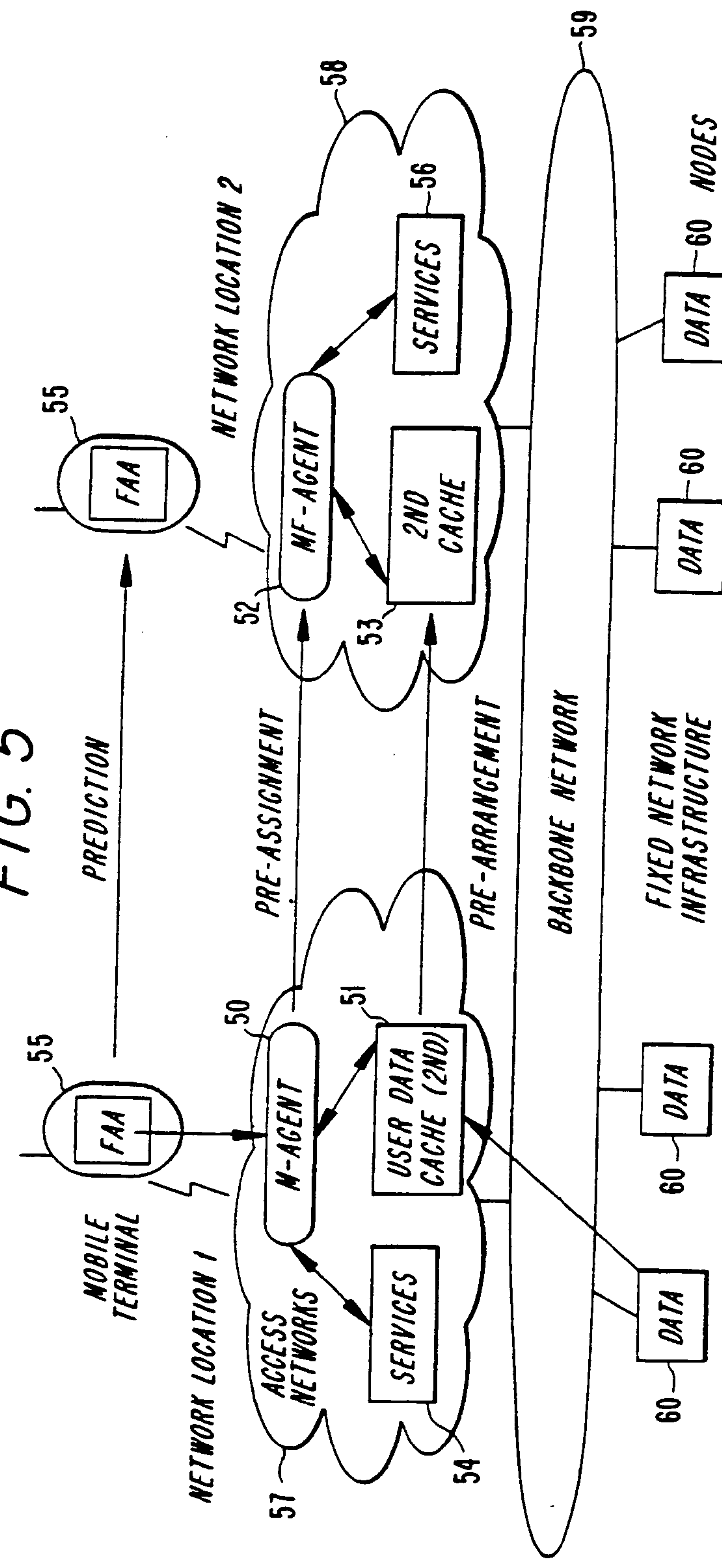
FIG. 1


USER
PERSONAL MOBILITY
USER
ID
TERMINAL
TERMINAL MOBILITY
10
12
ACCESS NETWORK
SERVICE/RESOURCE MOBILITY
SERVICE DATA
SERVICE DATA
14
BACKBONE NETWORK
SERVICE/RESOURCE/DATA MOBILITY
16
16
DATA MOBILITY
16
16
DATA NODES
DATA NODES
APPLICATION NODES



41
MOBILE FILE MANAGER WITH LCPM
42
MOBILE DATABASE MANAGER WITH PCP
43
WINDOWING APPLICATION SUPPORT
44
OTHER APPLICATIONS
45
MOBILE DISTRIBUTED SYSTEM PLATFORM (MDSP)
LOCATION SENSITIVE INFORMATION MANAGEMENT (LSIM)
47
PREDICTIVE MOBILITY MANAGEMENT FUNCTION (PMM)
46
31

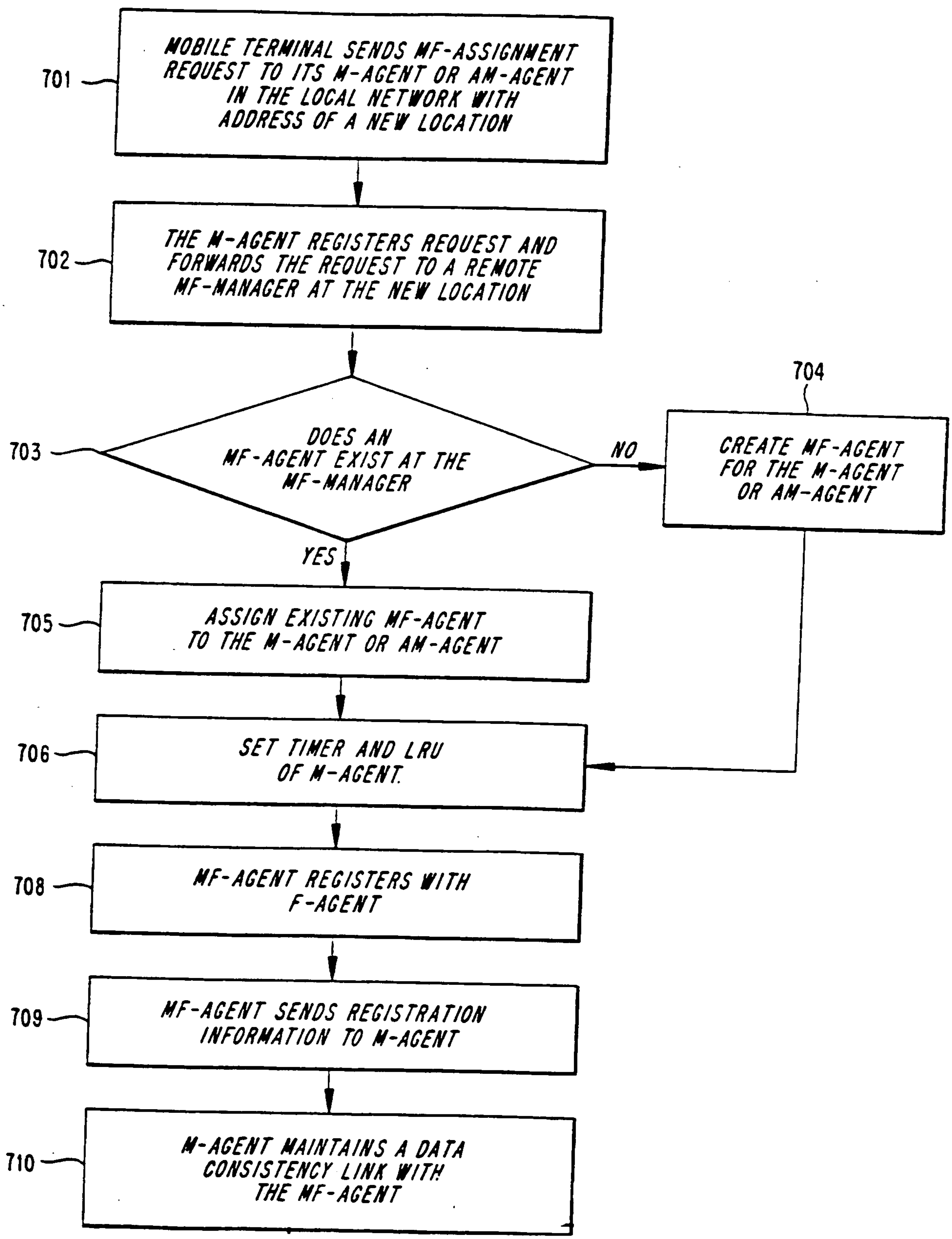
FIG. 4



55
FAA
MOBILE TERMINAL
PREDICTION
55
FAA
50
M-AGENT
51
USER DATA CACHE (2ND)
SERVICES
54
ACCESS NETWORKS
NETWORK LOCATION 1
57
PRE-ASSIGNMENT
PRE-ARRANGEMENT
NETWORK LOCATION 2
52
MF-AGENT
53
2ND CACHE
SERVICES
56
58
BACKBONE NETWORK
FIXED NETWORK INFRASTRUCTURE
59
DATA
60
NODES

FIG. 5

FIG. 7


701
MOBILE TERMINAL SENDS MF-ASSIGNMENT REQUEST TO ITS M-AGENT OR AM-AGENT IN THE LOCAL NETWORK WITH ADDRESS OF A NEW LOCATION
702
THE M-AGENT REGISTERS REQUEST AND FORWARDS THE REQUEST TO A REMOTE MF-MANAGER AT THE NEW LOCATION
703
DOES AN MF-AGENT EXIST AT THE MF-MANAGER
NO
704
CREATE MF-AGENT FOR THE M-AGENT OR AM-AGENT
YES
705
ASSIGN EXISTING MF-AGENT TO THE M-AGENT OR AM-AGENT
706
SET TIMER AND LRU OF M-AGENT.
708
MF-AGENT REGISTERS WITH F-AGENT
709
MF-AGENT SENDS REGISTRATION INFORMATION TO M-AGENT
710
M-AGENT MAINTAINS A DATA CONSISTENCY LINK WITH THE MF-AGENT

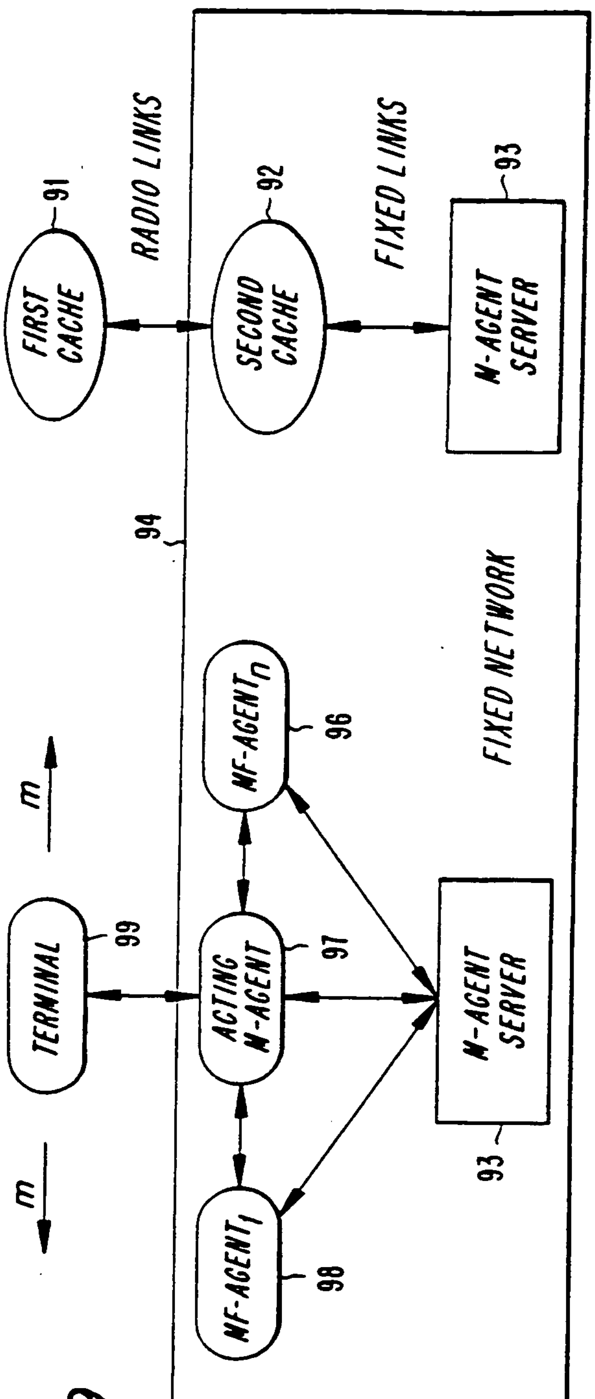

FIRST CACHE
91
RADIO LINKS
SECOND CACHE
92
FIXED LINKS
M-AGENT SERVER
93
94
TERMINAL
99
ACTING M-AGENT
97
NF-AGENT$_n$
96
NF-AGENT$_1$
98
M-AGENT SERVER
93
FIXED NETWORK
m
m

FIG. 9

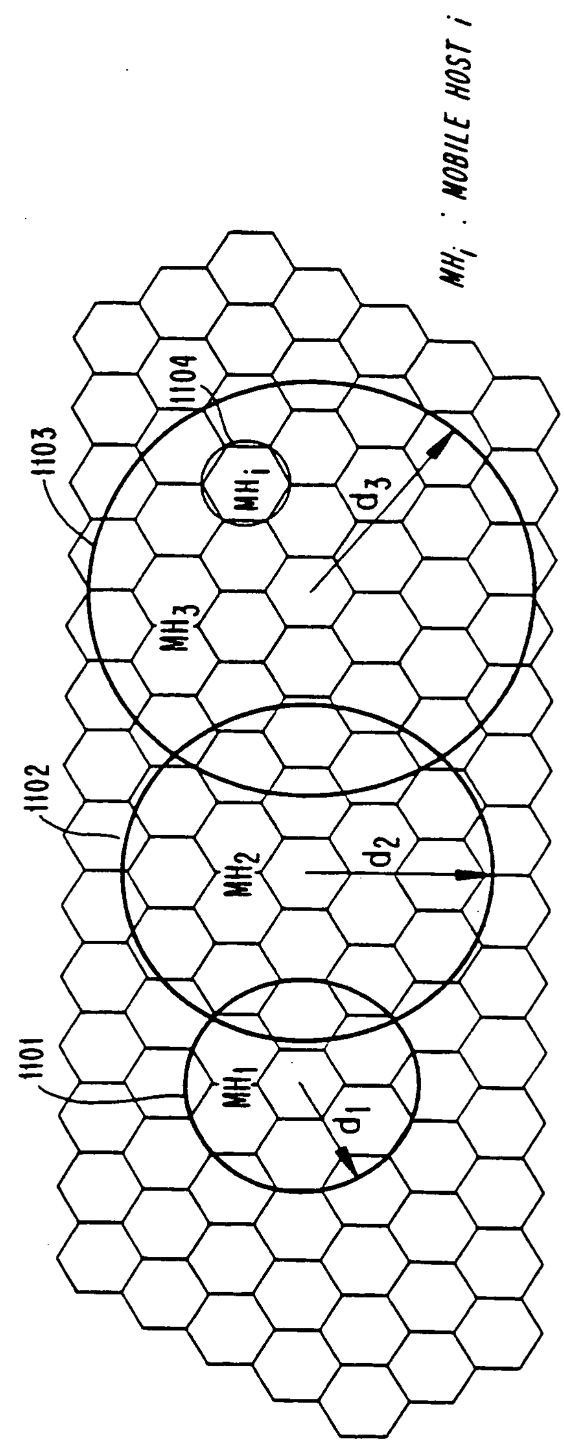

MH$_i$: MOBILE HOST i
1104
1103
MH$_i$
d$_3$
MH$_3$
1102
MH$_2$
d$_2$
1101
MH$_1$
d$_1$

FIG. 11

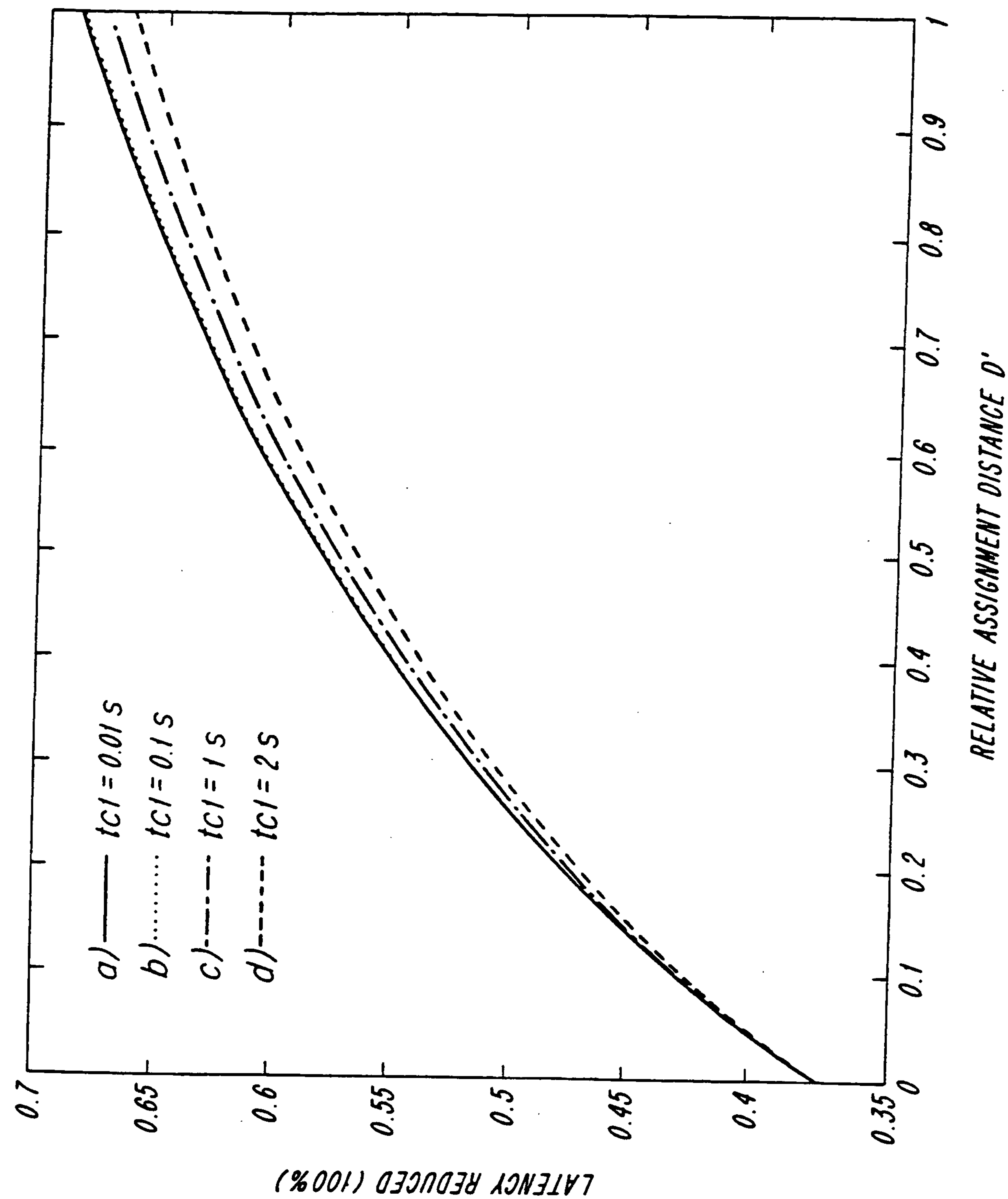

a) —— tc1 = 0.01 s
b) ········ tc1 = 0.1 s
c) -·-·- tc1 = 1 s
d) - - - - tc1 = 2 s
LATENCY REDUCED (100%)
RELATIVE ASSIGNMENT DISTANCE D'
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

FIG. 13

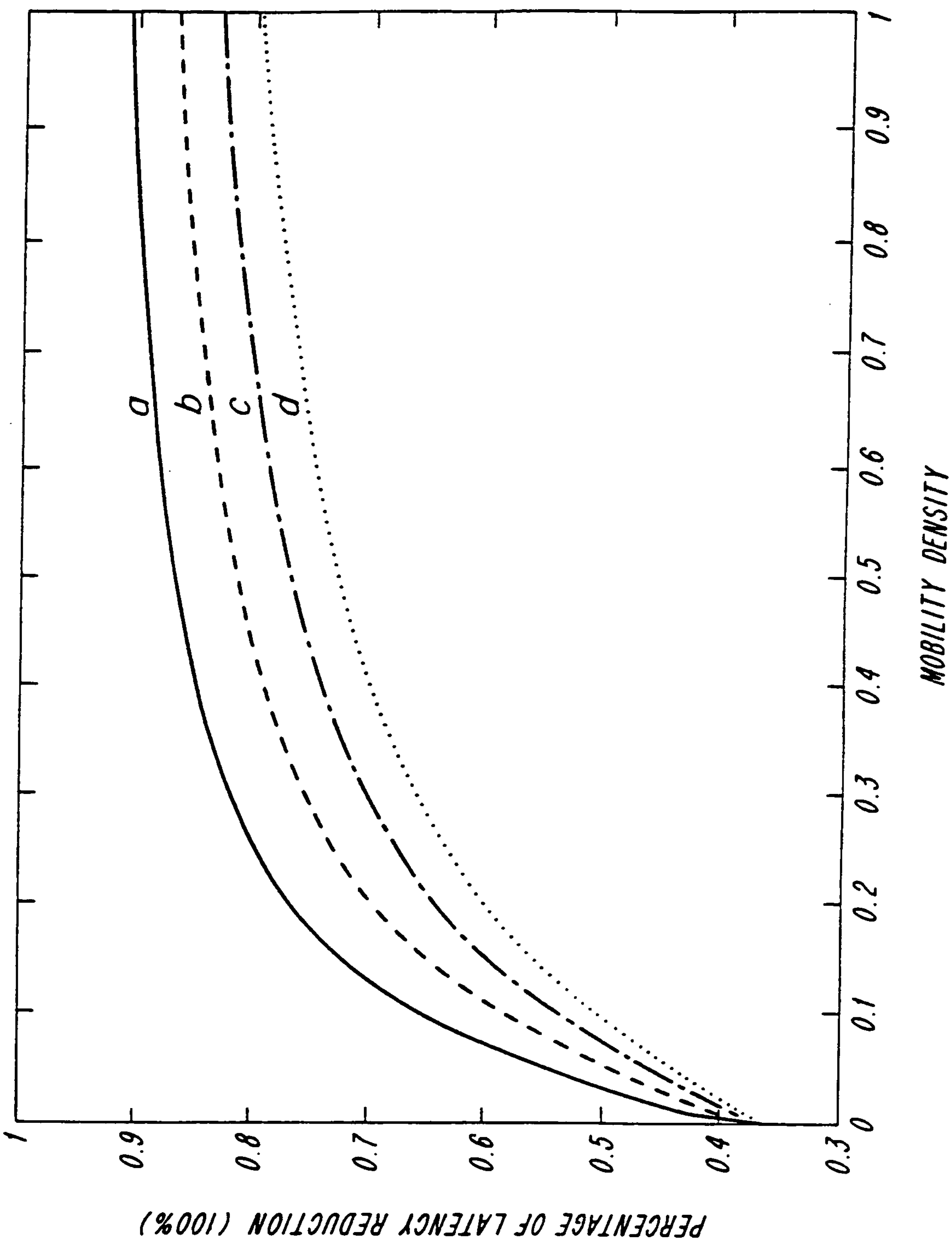

a
b
c
d
MOBILITY DENSITY
PERCENTAGE OF LATENCY REDUCTION (100%)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0.30
0.4
0.5
0.6
0.7
0.8
0.9
1

FIG. 15

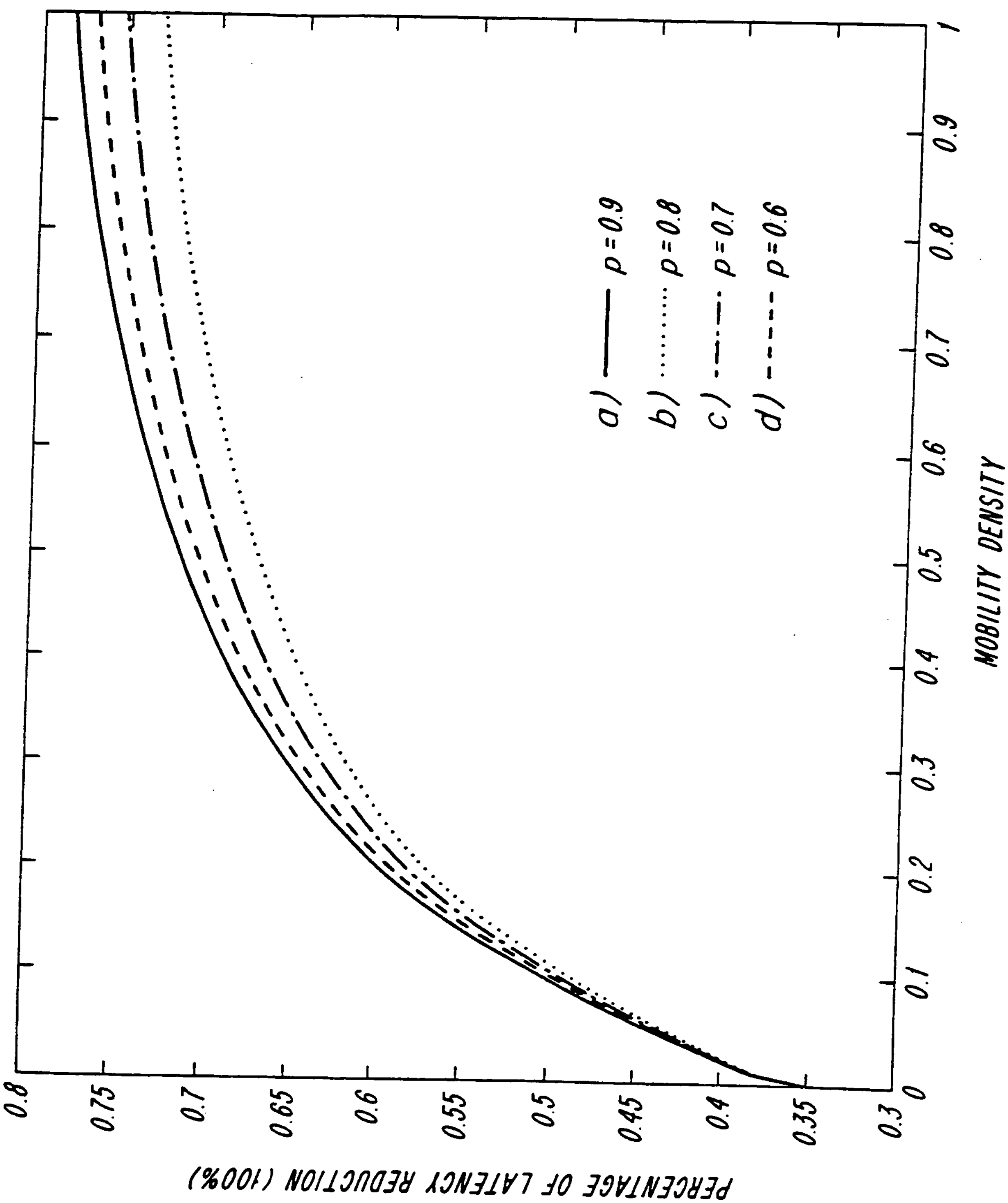

a) p=0.9
b) p=0.8
c) p=0.7
d) p=0.6
MOBILITY DENSITY
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
PERCENTAGE OF LATENCY REDUCTION (100%)
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8

FIG. 17

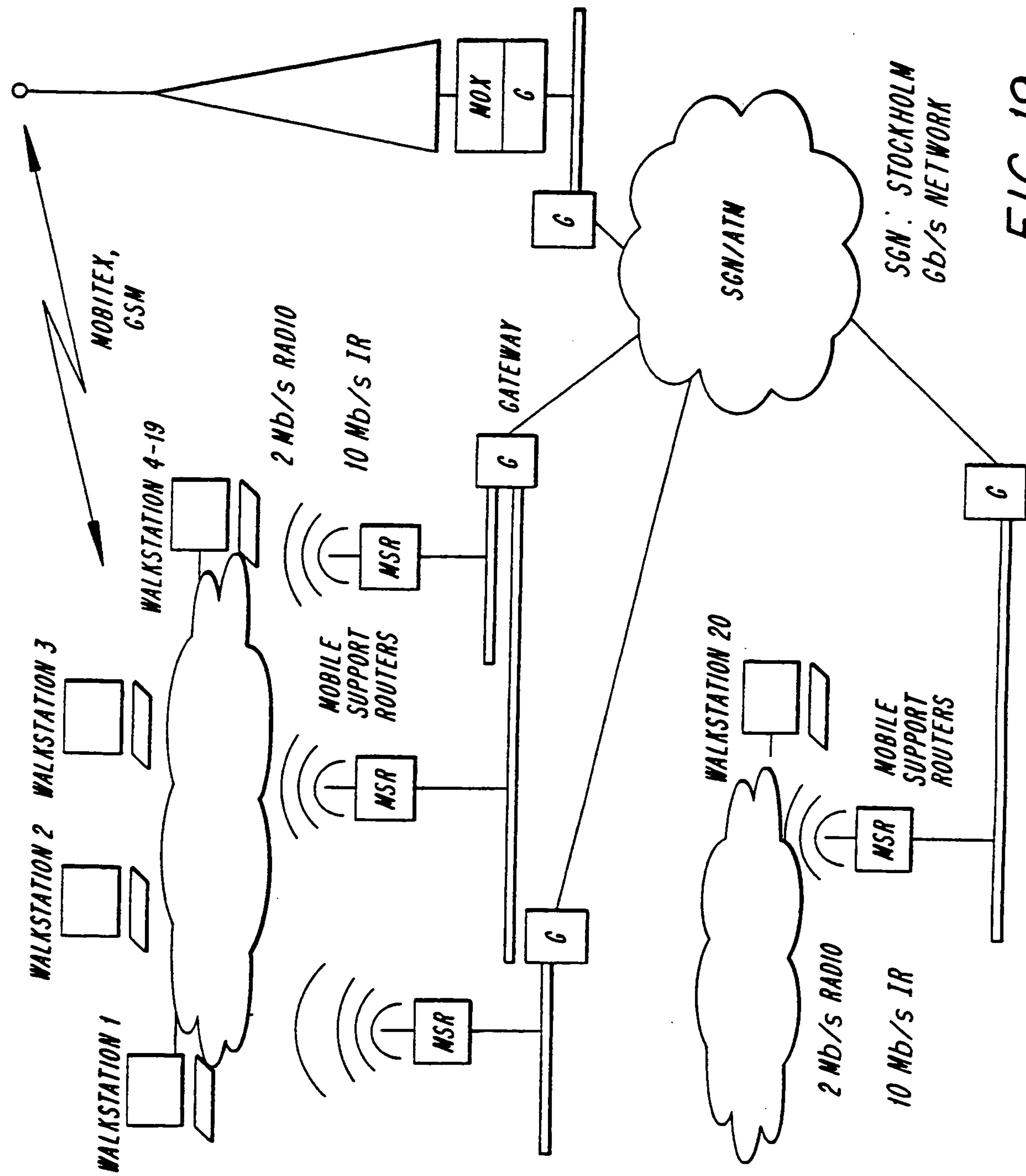
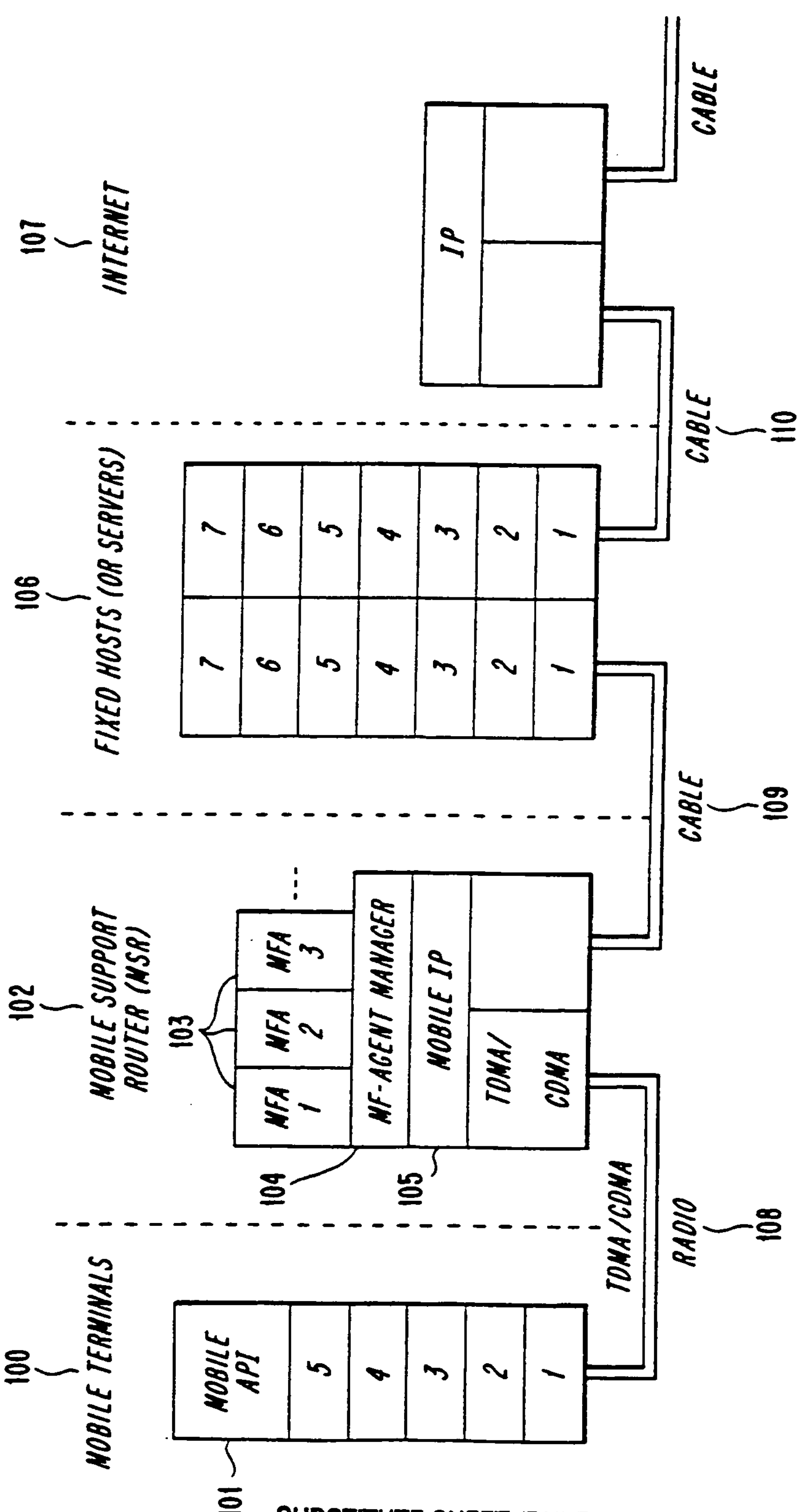

MOX
G
G
MOBITEX,
GSM
SGN/ATM
SGN : STOCKHOLM
Gb/s NETWORK
2 Mb/s RADIO
10 Mb/s IR
GATEWAY
G
WALKSTATION 4-19
MSR
MOBILE
SUPPORT
ROUTERS
G
WALKSTATION 3
WALKSTATION 2
WALKSTATION 1
MSR
MSR
G
WALKSTATION 20
MOBILE
SUPPORT
ROUTERS
MSR
2 Mb/s RADIO
10 Mb/s IR

FIG. 19

FIG. 21



SUBSTITUTE SHEET (RULE 26)

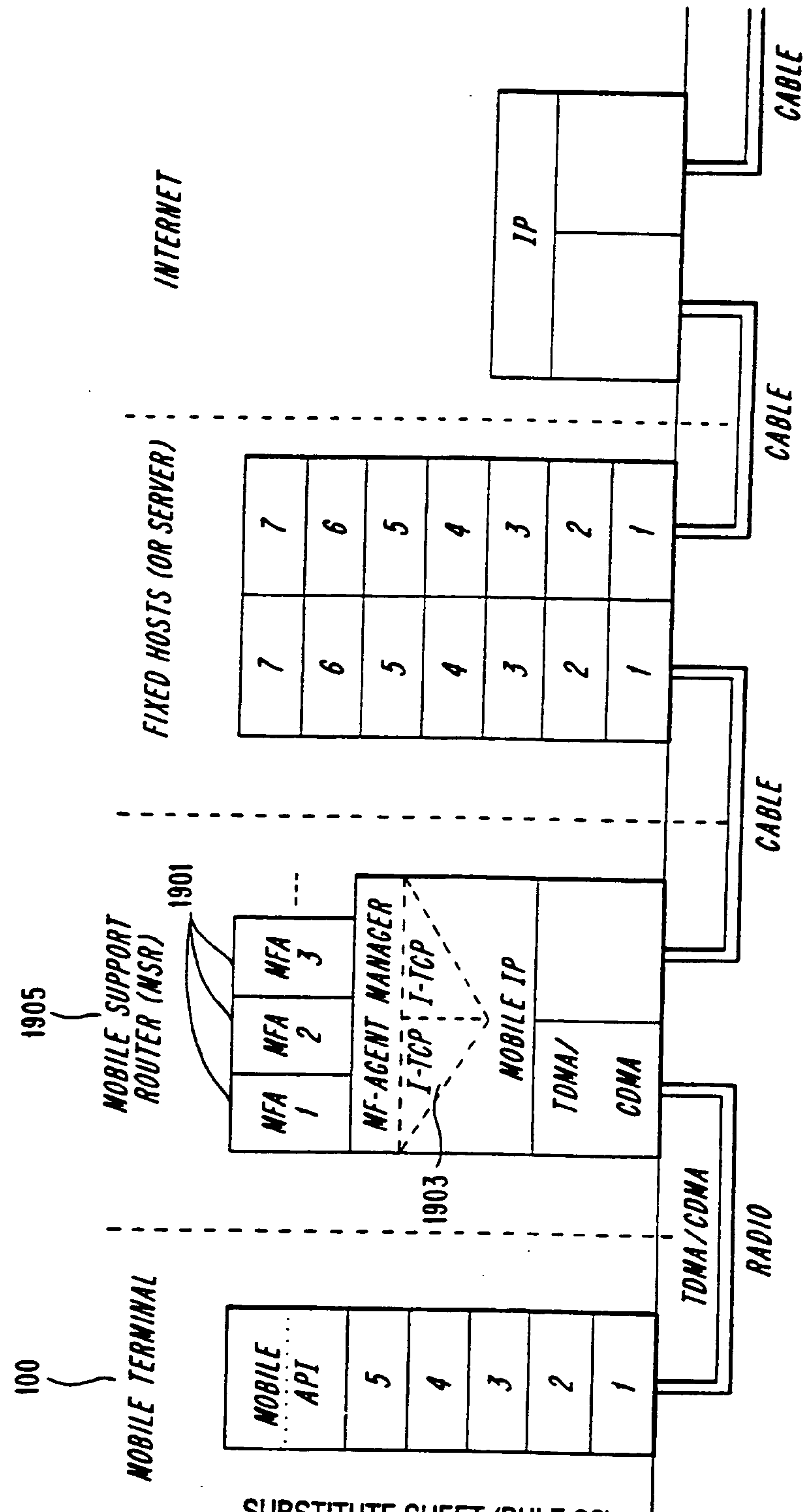

INTERNET
IP
CABLE
CABLE
CABLE
FIXED HOSTS (OR SERVER)
7
6
5
4
3
2
1
7
6
5
4
3
2
1
1905
MOBILE SUPPORT ROUTER (MSR)
1901
MFA 1
MFA 2
MFA 3
MF-AGENT MANAGER
I-TCP
I-TCP
MOBILE IP
TDMA/ CDMA
1903
TDMA/CDMA
RADIO
100
MOBILE TERMINAL
MOBILE API
5
4
3
2
1

FIG. 23

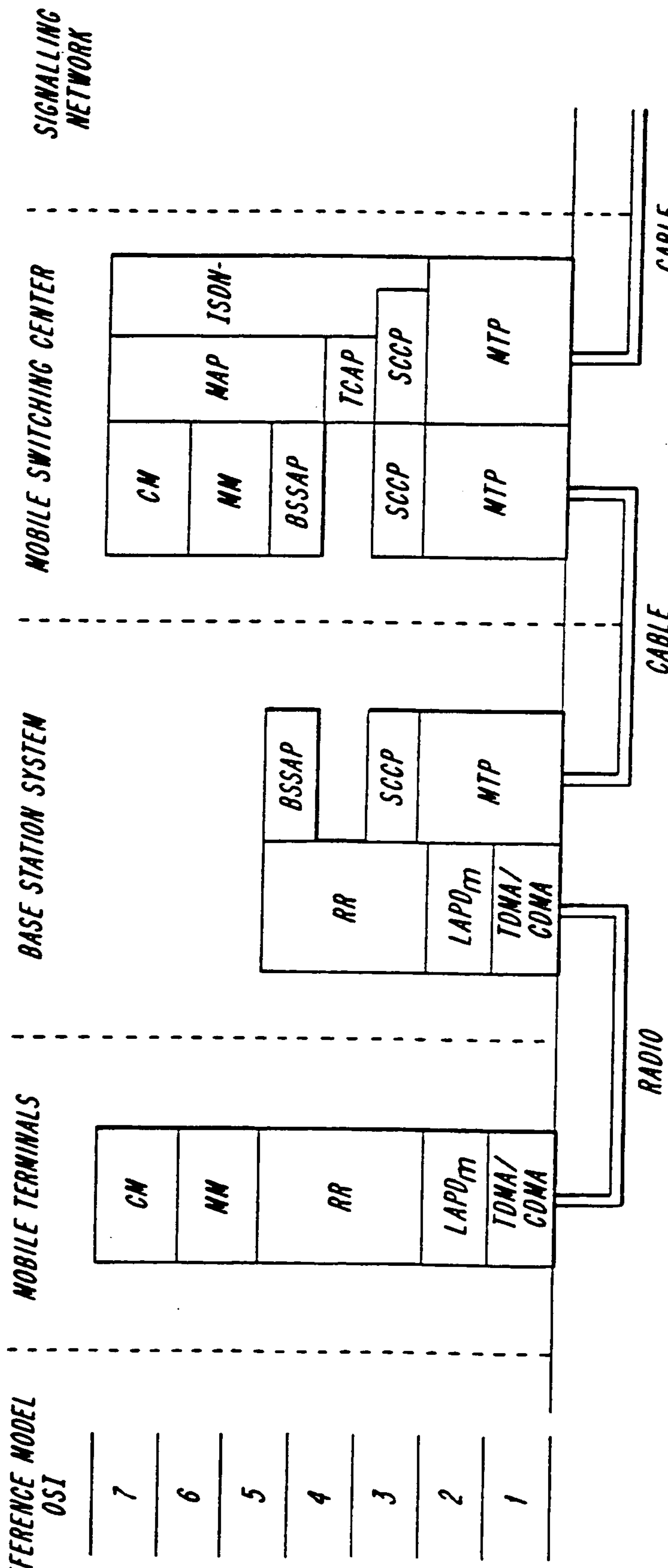


FIG. 25A
PRIOR ART

INTERNATIONAL SEARCH REPORT

International application No
PCT/SE 96/00905

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 H04Q7/24 H04Q3/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 H04Q H04M

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ERICSSON REVIEW, 1993, SWEDEN, vol. 70, no. 4, 1993, ISSN 0014-0171, pages 156-171, XP000415352 SODERBERG L: "Evolving an intelligent architecture for personal telecommunication" cited in the application see page 159, right-hand column, line 25 - page 160, right-hand column, line 24 see page 161, right-hand column, line 17 - page 169, right-hand column, line 2; figures ---	1,15-18, 22,26,28

-/--

[X] Further documents are listed in the continuation of box C.

[ ] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
16 December 1996	30. 12. 96

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+ 31-70) 340-3016	Janyszek, J-M

International application No
PCT/SE 96/00905

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	1994 ACM SIGMOD INTERNATIONAL CONFERENCE ON MANAGEMENT OF DATA, MINNEAPOLIS, MN, USA, 24-27 MAY 1994, vol. 23, no. 2, 24 - 27 May 1994, ISSN 0163-5808, SIGMOD RECORD, JUNE 1994, USA, pages 1-12, XP000603207 BARBARA D ET AL: "Sleepers and workaholics: Caching strategies in mobile environments" cited in the application see paragraph 7 -----	1,15-18, 22,26,28

1